hw17

```
library(MASS)
```

Question 1

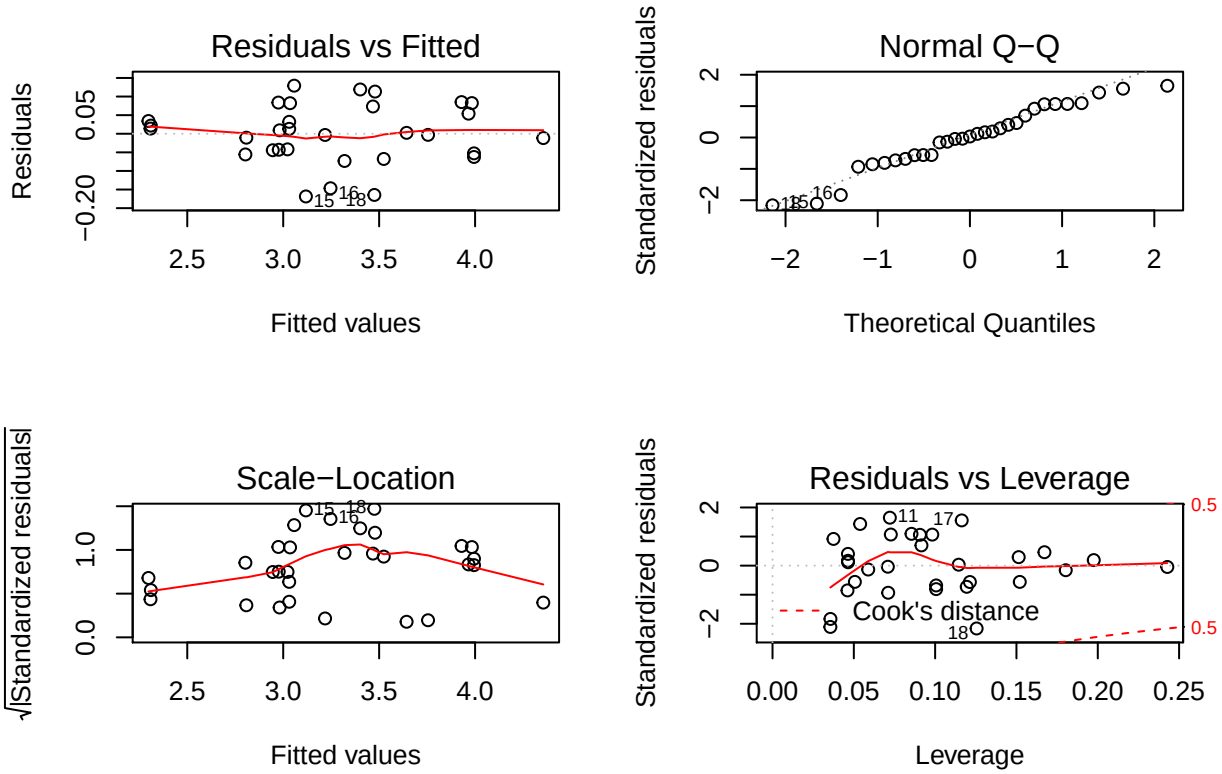
We linearlize our model by taking the *log* of the variables. This would get rid of any exponents and turn them into standard coefficients.

```
attach(trees, warn.conflicts = F)
V_log <- lm(log(Volume)~log(Girth)+log(Height))
summary(V_log)
```

```
##
## Call:
## lm(formula = log(Volume) ~ log(Girth) + log(Height))
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.168561 -0.048488  0.002431  0.063637  0.129223
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -6.63162    0.79979  -8.292 5.06e-09 ***
## log(Girth)   1.98265    0.07501  26.432  < 2e-16 ***
## log(Height)  1.11712    0.20444   5.464 7.81e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.08139 on 28 degrees of freedom
## Multiple R-squared:  0.9777, Adjusted R-squared:  0.9761
## F-statistic: 613.2 on 2 and 28 DF,  p-value: < 2.2e-16
```

Question 2

```
par(mfrow=c(2,2))
plot(V_log)
```



Looking at our diagnositc plots, we can see that our Residuals Vs Fitted plot is very nicely formed with equal vertical spread between points and a flat trend line. Our Normal Q-Q plot also has a linear shape, although some breaks are shown. These breaks probably occur from our linearization, but the overall shape shows good normal distribution in our residuals. The scale location plot is slightly concerning with its trendline not being flat and higher spread on the residuals near the center tather than random, indicating that we might be overfitting. Finally, our Residuals vs Leverage looks good, indicating that most of our data is decently fit with the usual outliers. Overall, these diagnostic plots show that our linearization model is a good fit for out data.

Question 3

The summary from question 1 tells us that $\beta_1 = 1.98265$ and $\beta_2 = 1.11712$

```
confint(V_log, level = 0.95)
```

```
##                 2.5 %    97.5 %
## (Intercept) -8.269912 -4.993322
## log(Girth)   1.828998  2.136302
## log(Height)  0.698353  1.535894
```

Seeing the confidence interval for β_1 is from 1.828 to 2.136, we can see that our β_1 is within that. In addition, our β_2 is within the confidence interval of 0.698 to 1.535, so the CI contains the theoretical values suggested by the relationship.

Question 4

```
newdata = data.frame(Girth=10.9, Height=75)
predict(V_log, newdata, interval = "predict")
```

```
##       fit     lwr    upr
## 1 2.92763 2.75656 3.0987
```

Question 5

```
exp(2.75656)
```

```
## [1] 15.74558
```

```
exp(3.0987)
```

```
## [1] 22.16911
```